SYMPLECTIC RECONSTRUCTION OF DATA FOR HEAT AND WAVE EQUATIONS

JESPER CARLSSON

ABSTRACT. This report concerns the inverse problem of estimating a spacially dependent coefficient of a partial differential equation from observations of the solution at the boundary. Such a problem can be formulated as an optimal control problem with the coefficient as the control variable and the solution as state variable. The heat or the wave equation is here considered as state equation. It is well known that such inverse problems are ill-posed and need to be regularized. The powerful Hamilton-Jacobi theory is used to construct a simple and general method where the first step is to analytically regularize the Hamiltonian; next its Hamiltonian system, a system of nonlinear partial differential equations, is solved with the Newton method and a sparse Jacobian.

CONTENTS

1. INTRODUCTION

In this paper we study the inverse problem to determine a spacially dependent coefficient σ of a partial differential equation from partial knowledge of the forward solution u. In particular, we seek the diffusion coefficient in the heat equation and the wave speed coefficient in the wave equation. Inverse problems arise in many applications such as inverse scattering, impedance tomography and topology optimization, see *e.g.* [1, 3, 6, 14], and share the property that they are ill posed *i.e.* given data u there may not exist a corresponding coefficient σ, and if it exists it may not be unique nor depend continuously on u. To be able to determine σ

2000 *Mathematics Subject Classification.* Primary: 65M32; Secondary: 49L25.

Key words and phrases. Inverse Problems, Parameter Reconstruction, Hamilton-Jacobi, Regularization.

Support by the Swedish Research Council grants 2002-6285 and 2002-4961, and the European network HYKE, funded by the EC as contract HPRN-CT-2002-00282, is acknowledged.

the problem thus needs to be regularized such that it becomes well posed. The method used here to regularize and to solve the inverse problem is based on the work [7, 8, 15, 16] where the inverse problem is formulated as an optimal control problem and the corresponding Hamilton-Jacobi equation is used to construct a regularization, to obtain convergence results, and to finally solve the regularized problem by using the method of characteristics, *i.e.* to solve the corresponding Hamiltonian system.

The paper is stuctured as follows: In Section 2 the general theory of optimal control of partial differential equations and Hamilton-Jacobi-Bellman is presented. In Section 3 the idea of how to optimally control the heat equation is discussed together with numerical examples, and in Section 4 the control of the wave equation is treated.

2. Optimal Control and Dynamic Programming

Consider a differential equation constrained minimization problem with solution $\varphi : \Omega \times [0,T] \to \mathbb{R}$, $\varphi(\cdot,t) \in V$ and control $\sigma : \Omega \times [0,T] \to B$, $\sigma(\cdot,t) \in W$ for an open domain Ω, some Hilbert spaces V and W on Ω, and a closed bounded set $B \subset \mathbb{R}$:

$$\begin{aligned} &\min_{\sigma:\Omega\times[0,T]\to B} \int_0^T h(\varphi,\sigma)\ \mathrm{d}t + g(\varphi^T), \\ &\varphi_t = f(\varphi,\sigma), \end{aligned} \tag{1}$$

with $\varphi^T := \varphi(\cdot,T)$ and given initial value $\varphi^0 = \varphi(\cdot,0)$. Here, φ_t denotes the partial derivative with respect to time, $f : V \times W \to V$ is the flux, and $h : V \times W \to \mathbb{R}$, $g : V \to \mathbb{R}$ are given functions.

This optimal control problem can be solved either directly using constrained minimization or by dynamic programming. The Lagrangian becomes

$$L(\varphi,\lambda,\sigma) := \int_0^T \langle \lambda, f(\varphi,\sigma) - \varphi_t \rangle + h(\varphi,\sigma)\ \mathrm{d}t,$$

with Lagrange multiplier $\lambda : \Omega \times [0,T] \to \mathbb{R}$, $\lambda(\cdot,t) \in V$, and the constrained minimization method is based on the Pontryagin method

$$\begin{aligned} \varphi_t &= f(\varphi,\sigma), \\ \lambda_t &= -\langle \lambda, f_\varphi(\varphi,\sigma)\rangle + h_\varphi(\varphi,\sigma), \\ \sigma(\cdot,t) &\in \operatorname*{argmin}_{a:\Omega\to B} \{\langle \lambda, f(\varphi,a)\rangle + h(\varphi,a)\}. \end{aligned} \tag{2}$$

with given initial value φ^0, final value $\lambda^T := \lambda(\cdot,T) = g_\varphi(\varphi^T)$, and where f_φ, h_φ denotes the Gateaux derivatives with respect to φ and $\langle v,w\rangle$ is the duality pairing on V, which reduces to the $L^2(\Omega)$ inner product if $v,w \in L^2(\Omega)$. For a differentiable Lagrangian that is convex in σ the Pontryagin principle coincides with the Lagrangian formulation for a constrained interior minimum

$$\begin{aligned} \varphi_t &= f(\varphi,\sigma), \\ \lambda_t &= -\langle \lambda, f_\varphi(\varphi,\sigma)\rangle + h_\varphi(\varphi,\sigma) \\ 0 &= \langle \lambda, f_\sigma(\varphi,\sigma)\rangle + h_\sigma(\varphi,\sigma), \\ \sigma &\in B, \end{aligned} \tag{3}$$

but in general (2) and (3) may have different solutions φ,λ,σ although both describe necessary conditions for a minimizer to (1). If an explicit minimizer in (2) can be found the Pontryagin principle gives additional information about the control.

Pontryagin's minimum principle can also be written as a Hamiltonian system, see [2],

$$\begin{aligned} \varphi_t &= H_\lambda(\varphi, \lambda) \\ \lambda_t &= -H_\varphi(\varphi, \lambda) \end{aligned} \tag{4}$$

with φ^0 given, $\lambda^T = g_\varphi(\varphi^T)$, and the Hamiltonian $H : V \times V \to \mathbb{R}$ defined as

$$H(\lambda, \varphi) \quad := \min_{a:\Omega\to B} \{\langle \lambda, f(\varphi, a)\rangle + h(\varphi, a)\}. \tag{5}$$

The alternative dynamic programming method is based on the value function $U : V \times [0, T] \to \mathbb{R}$,

$$U(\phi, \tau) := \inf_{\sigma:\Omega\times[\tau,T]\to B} \left\{ \int_\tau^T h(\varphi, \sigma)\ \mathrm{d}t + g(\varphi^T) \;\middle|\; \varphi_t = f(\varphi, \sigma),\ \varphi(\cdot, \tau) = \phi \in V \right\}$$

which solves the nonlinear Hamilton-Jacobi-Bellman equation

$$\partial_t U(\phi, t) + H\big(U_\phi(\phi, t), \phi\big) = 0, \quad U(\phi, T) = g(\phi), \tag{6}$$

with Hamiltonian defined as in (5). Note that solving the Hamiltonian system (4) is the method of characteristics for the Hamilton-Jacobi equation (6), with $\lambda(x, t) = U_\varphi(\varphi(x, t), t)$. In general, the value function is however not everywhere differentiable and the multiplier λ becomes ill defined in a classical sense.

The Hamilton-Jacobi formulation (6) has the advantages that there is a complete well-posedness theory for Hamilton-Jacobi equations, based on non-differential viscosity solutions, see [9], and it finds a global minimum. However, (6) is not computationally feasible for problems in high dimension, such as the case where φ is an approximation of a solution to a partial differential equation. The Hamiltonian form (4) has the advantage that it is computationally feasible but the drawbacks are that it only focuses on local minima and that the Hamiltonian (5) in general only is Lipschitz continuous, even if f, g and h are smooth, which means that the optimal control depends discontinuously on (λ, φ) and (4) becomes undefined where the Hamiltonian is not differentiable.

In the following sections we will use a regularized version of (4) to iteratively solve the nonlinear constrained optimization problem (1).

3. Parameter Reconstruction for the Heat Equation

A distributed parameter reconstuction problem for the heat equation is to find a heat conductivity (the control) *e.g.* $\sigma : \bar{\Omega} \times [0, T] \to [\sigma_-, \sigma_+]$, $\sigma(\cdot, t) \in W$, $0 < \sigma_- < \sigma_+$, and a temperature distribution (the state) $u : \bar{\Omega} \times [0, T] \to \mathbb{R}$, $u(\cdot, t) \in V$ that satifies the heat equation

$$\begin{aligned} u_t &= \operatorname{div}(\sigma \nabla u), && \text{in } \Omega \times (0, T], \\ \sigma \nabla u \cdot \mathbf{n} &= j, && \text{on } \partial\Omega \times (0, T], \\ u &= 0, && \text{on } \bar{\Omega} \times \{t = 0\}, \end{aligned} \tag{7}$$

such that the error functional

$$\int_0^T \int_{\partial\Omega} (u - u^*)^2\ \mathrm{d}s\ \mathrm{d}t, \tag{8}$$

is minimized. The function $u^* = u^*(x, t)$ often represents physical measurements contaminated by some noise, *e.g.* $u^*(x, t) = u_{true}(x, t) + w(x, t)$ where w is a noise term and u_{true} satisfies the above heat equation for some unknown parameter σ_{true}, and in practice the control is only spacially dependent, $\sigma_{true} = \sigma_{true}(x)$. The primary goal is thus to determine the unknown diffusion coefficient σ_{true} and the method to do so is to minimize the objective functional (8).

Inverse problems like (7), (8) are in general ill-posed due to one or more of the following reasons:

(1) There exists no minimizer (u, σ), something that may occur with noisy data. Given unperturbed data u^* corresponding to σ_{true}, it is evident that there exists a minimizer to (7), (8).
(2) The minimizer is not unique, *e.g.* although it may be possible to find an optimal state that minimizes (8), u and σ may not be unique in Ω.
(3) The solution (u, σ), and particularly the control σ, depends discontinuously on data u^*.

A simple and common way to impose well-posedness to many inverse problems is to add a Tikhonov regularization of the form $\epsilon\|\sigma\|^2_{L^2(\Omega\times(0,T))}$ for $\epsilon > 0$, to the objective functional (8), see [1, 10, 14, 17]. Using the Pontryagin principle presented in the previous section we will in Section 3.2 regularize the inverse problem (7), (8) in a way that is comparable to a Tikhonov regularization.

Formulated as an optimal control problem the most natural assumption on the control σ is that it is dependent on both time and space but as we will see in Section 3.3 it is also possible to let $\sigma = \sigma(x)$, $\sigma = \sigma(t)$, or even let σ be constant in time and space.

3.1. **The Hamiltonian System.** Following Section 2 the Hamiltonian associated to the optimal control problem (7) and (8) is

$$\begin{aligned} H(u,q,t) &:= \min_{\sigma:\Omega\to[\sigma_-,\sigma_+]} \int_{\partial\Omega} (u-u^*)^2 \,\mathrm{d}s + \int_\Omega \operatorname{div}(\sigma\nabla u) q \,\mathrm{d}x \\ &= \int_{\partial\Omega} (u-u^*)^2 + jq \,\mathrm{d}s + \min_{\sigma:\Omega\to[\sigma_-,\sigma_+]} \int_\Omega -\sigma\nabla u\cdot\nabla q \,\mathrm{d}x \\ &= \int_{\partial\Omega} (u-u^*)^2 + jq \,\mathrm{d}s - \int_\Omega \underbrace{\max_{\sigma\in[\sigma_-,\sigma_+]} \{\sigma\nabla u\cdot\nabla q\}}_{\mathfrak{h}(\nabla u\cdot\nabla q)} \,\mathrm{d}x. \end{aligned} \tag{9}$$

and the Hamiltonian system, in strong form, then becomes

$$\begin{aligned} u_t &= \operatorname{div}\big(\tilde\sigma\nabla u\big), && \text{in } \Omega\times(0,T], \\ \tilde\sigma\nabla u\cdot\mathbf{n} &= j, && \text{on } \partial\Omega\times(0,T], \\ u &= 0, && \text{on } \bar\Omega\times\{t=0\}, \\ -q_t &= \operatorname{div}\big(\tilde\sigma\nabla q\big), && \text{in } \Omega\times(0,T], \\ \tilde\sigma\nabla q\cdot\mathbf{n} &= 2(u-u^*), && \text{on } \partial\Omega\times(0,T], \\ q &= 0, && \text{on } \Omega\times\{t=T\}, \end{aligned} \tag{10}$$

with

$$\tilde\sigma := \mathfrak{h}'(\nabla u\cdot\nabla q). \tag{11}$$

It is here evident that the Hamiltonian only is Lipschitz continuous and the control $\tilde\sigma$ is a bang-bang type control which depends discontinuously on the solutions (u, q), see Figure 1. From the optimality conditions (3) an optimal solution has to satisfy $\nabla u\cdot\nabla q = 0$ and (10) is thus undefined since $\mathfrak{h}'(0)$ is set valued, which calls for a regularization.

3.2. **Regularization.** A simple regularization of the Hamiltonian system (10), and consequently of the Hamiltonian (9), is to approximate $\mathfrak{h}'$ with the parabolic function

$$\mathfrak{h}'_\delta(\nabla u \cdot \nabla q) := \underbrace{\frac{\sigma_+ + \sigma_-}{2}}_{\bar{\sigma}} + \underbrace{\frac{\sigma_+ - \sigma_-}{2}}_{\hat{\sigma}} \tanh(\frac{1}{\delta}\nabla u \cdot \nabla q), \tag{12}$$

for some small $\delta > 0$, see Figure 1. This regularization can be compared with a classic Tikhonov regularization where a small L^2-penalty of the control is added to the objective function (8), *i.e.* to minimize

$$\int_0^T \int_{\partial\Omega} (u - u^*)^2 \, \mathrm{d}s \, \mathrm{d}t + \delta \int_0^T \int_\Omega \sigma^2 \, \mathrm{d}x \, \mathrm{d}t. \tag{13}$$

Minimizing (13) under the constraint (7) will lead to a C^2-Hamiltonian with

$$H(u, q, t) = \int_{\partial\Omega} (u - u^*)^2 + jq \, \mathrm{d}s - \int_\Omega \underbrace{\max_{\sigma \in [\sigma_-, \sigma_+]} \{\sigma(\nabla u \cdot \nabla q - \delta\sigma)\}}_{\mathfrak{h}_{Tikhonov}(\nabla u \cdot \nabla q)} \, \mathrm{d}x,$$

which can be seen in Figure 1.

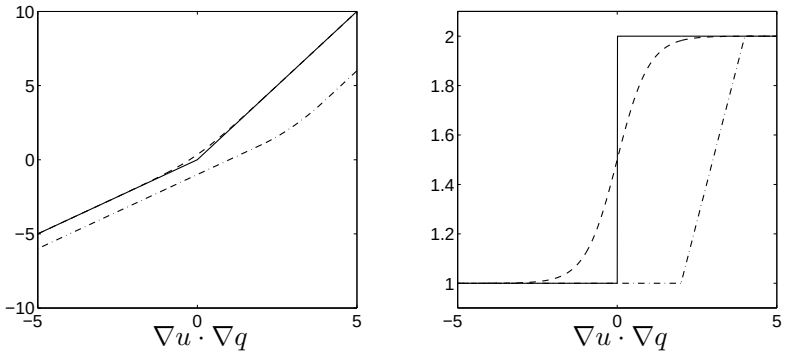

FIGURE 1. The functions $\mathfrak{h}$ (solid line), $\mathfrak{h}_\delta$ (dashed line), $\mathfrak{h}_{Tikhonov}$ (dash-dotted line) to the left and their derivatives to the right.

Another way to describe the simple regularization (12) is to see what kind of penalty on the objective function it corresponds to. We note that the regularized Hamiltionian system can be written as

$$\int_\Omega -u_t v - \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)\nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} jv \, \mathrm{d}s = 0, \quad \forall v \in V,$$
$$\int_\Omega q_t v - \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)\nabla q \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} 2(u - u^*)v \, \mathrm{d}s = 0, \quad \forall v \in V,$$

or by a redefinition of σ

$$\begin{aligned} \int_\Omega -u_t v - \sigma\nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} jv \, \mathrm{d}s = 0, \quad & \forall v \in V, \\ \int_\Omega q_t v - \sigma\nabla q \cdot \nabla v \, \mathrm{d}x + \int_{\partial\Omega} 2(u - u^*)v \, \mathrm{d}s = 0, \quad & \forall v \in V, \\ \int_\Omega \Big(\sigma - \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)\Big)v \, \mathrm{d}x = 0, \quad & \forall v \in W. \end{aligned} \tag{14}$$

Let $\mathfrak{H}$ be the primitive function of the inverse function of $\mathfrak{h}'_\delta$ *i.e.*

$$\mathfrak{H}(\sigma) := \frac{\delta}{2\hat{\sigma}}\bigg((\sigma-\sigma_-)\ln\Big(\frac{\sigma-\sigma_-}{\hat{\sigma}}\Big)+(\sigma_+-\sigma)\ln\Big(\frac{\sigma_+-\sigma}{\hat{\sigma}}\Big)\bigg),$$

then it is evident that (14) can be seen as the first order optimality conditions for the problem to minimize

$$\int_0^T\int_{\partial\Omega}(u-u^*)^2\,\mathrm{d}s\,\mathrm{d}t+\int_0^T\int_\Omega\mathfrak{H}(\sigma)\,\mathrm{d}x\,\mathrm{d}t,$$

under the constraint (7). In Figure 2, the function $\mathfrak{H}(\sigma)$ is compared with a Tikhonov regularization of the form $\delta(\sigma-\bar{\sigma})^2$.

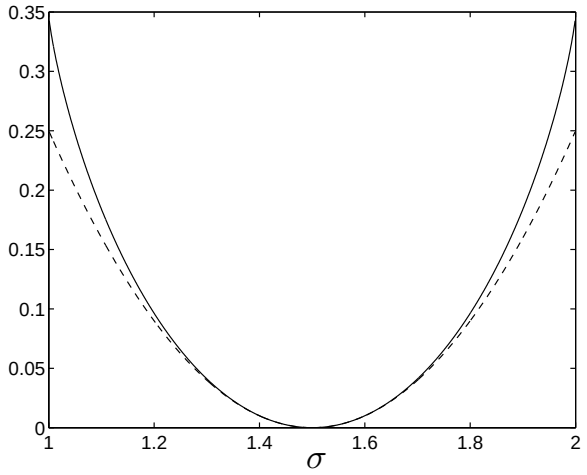

FIGURE 2. The function $\mathfrak{H}(\sigma)$ (solid line) compared to the L^2 penalty function $\delta(\sigma-\bar{\sigma})^2$ (dashed line) for $\delta=1$, $\sigma_-=1$ and $\sigma_+=2$.

It is often beneficial to prevent spacial oscillations of the coefficient by adding a penalty on the L^2-norm of the gradient of the coefficient, *i.e.* $\epsilon\|\nabla\sigma\|^2_{L^2(\Omega\times(0,T))}$, for $\epsilon>0$, to the objective function (8). For such a penalty the minimization in the corresponding Hamiltonian

$$H(u,q,t) := \min_{\sigma:\Omega\to[\sigma_-,\sigma_+]}\int_{\partial\Omega}(u-u^*)^2\,\mathrm{d}s+\int_\Omega \mathrm{div}(\sigma\nabla u)q+\epsilon|\nabla\sigma|^2\,\mathrm{d}x, \tag{15}$$

can not be done explicitly, and instead taking the first variation in σ would give the system

$$\begin{aligned} u_t &= \mathrm{div}\big(\sigma\nabla u\big),\\ -q_t &= \mathrm{div}\big(\sigma\nabla q\big),\\ 2\epsilon\Delta\sigma &= -\nabla u\cdot\nabla q,\\ \sigma &\in [\sigma_-,\sigma_+]. \end{aligned}$$

which corresponds to the usual first order optimality conditions for the Lagrangian. How to treat different penalties on the control in an optimal control setting is discussed in Section 3.4.

3.3. **Time Independent Control.** To study the case when the control σ is independent of time we first assume that it not only is independent of time but also depends on an auxilliary variable z, *i.e.* $\sigma : \bar{\Omega} \times [0, \tilde{T}] \to [\sigma_-, \sigma_+]$, $\sigma = \sigma(x, z)$. For a moment we also assume that $u : \bar{\Omega} \times [0, T] \times [0, \tilde{T}] \to \mathbb{R}$, $u = u(x, t, z)$, but with the same measurements as in (8). If we treat z as the time and t as a spacial variable we can define the optimal control problem

$$\min_{\sigma:\bar{\Omega}\times[0,\tilde{T}]\to[\sigma_-,\sigma_+]} \frac{1}{\tilde{T}} \int_0^{\tilde{T}} \int_0^T \int_{\partial\Omega} (u - u^*)^2 \, \mathrm{d}s \, \mathrm{d}t \, \mathrm{d}z, \tag{16}$$

where the state u satisfies the partial differential equation

$$\begin{aligned} u_z &= \frac{1}{\tilde{T}}\Big(\operatorname{div}(\sigma\nabla u) - u_t\Big), && \text{in } \Omega \times (0,T) \times (0,\tilde{T}], \\ \sigma\nabla u \cdot \mathbf{n} &= j, && \text{on } \partial\Omega \times (0,T) \times (0,\tilde{T}], \\ u &= 0, && \text{on } \bar{\Omega} \times \{t = 0\} \times (0,\tilde{T}], \\ u &= u_0, && \text{on } \bar{\Omega} \times (0,T) \times \{z = 0\}. \end{aligned} \tag{17}$$

for some arbitrary initial condition $u(x, t, 0) = u_0$.

The Hamiltonian for (16), (17) is

$$\begin{aligned} H(u,q,z) := &\min_{\sigma:\Omega\to[\sigma_-,\sigma_+]} \frac{1}{\tilde{T}} \int_0^T \int_{\partial\Omega} (u - u^*)^2 \, \mathrm{d}s \, \mathrm{d}t \\ &+ \frac{1}{\tilde{T}} \int_0^T \int_\Omega \Big(\operatorname{div}(\sigma\nabla u) - u_t\Big) q \, \mathrm{d}x \, \mathrm{d}t \\ =&\frac{1}{\tilde{T}} \int_0^T \int_{\partial\Omega} (u - u^*)^2 + jq \, \mathrm{d}s \, \mathrm{d}t - \frac{1}{\tilde{T}} \int_0^T \int_\Omega u_t q \, \mathrm{d}x \, \mathrm{d}t \\ &- \frac{1}{\tilde{T}} \int_\Omega \underbrace{\max_{\sigma\in[\sigma_-,\sigma_+]} \left\{ \sigma \int_0^T \nabla u \cdot \nabla q \, \mathrm{d}t \right\}}_{\mathfrak{h}\left(\int_0^T \nabla u \cdot \nabla q \, \mathrm{d}t\right)} \mathrm{d}x, \end{aligned} \tag{18}$$

and the Hamiltonian system is given by

$$\begin{aligned} u_z &= \frac{1}{\tilde{T}}\Big(\operatorname{div}(\mathfrak{h}'\nabla u) - u_t\Big), && \text{in } \Omega \times (0,T) \times (0,\tilde{T}], \\ \mathfrak{h}'\nabla u \cdot \mathbf{n} &= j, && \text{on } \partial\Omega \times (0,T) \times (0,\tilde{T}], \\ u &= 0, && \text{on } \bar{\Omega} \times \{t = 0\} \times (0,\tilde{T}], \\ u &= u_0, && \text{on } \bar{\Omega} \times (0,T) \times \{z = 0\}, \\ -q_z &= \frac{1}{\tilde{T}}\Big(\operatorname{div}(\mathfrak{h}'\nabla q) + q_t\Big), && \text{in } \Omega \times (0,T) \times (0,\tilde{T}], \\ \mathfrak{h}'\nabla q \cdot \mathbf{n} &= 2(u - u^*), && \text{on } \partial\Omega \times (0,T) \times (0,\tilde{T}], \\ q &= 0, && \text{on } \bar{\Omega} \times \{t = T\} \times (0,\tilde{T}], \\ q &= 0, && \text{on } \bar{\Omega} \times (0,T) \times \{z = \tilde{T}\}. \end{aligned} \tag{19}$$

Under the assumption that the solutions u and q in (19) are asymptotically stationary as $\tilde{T} \to \infty$, the Hamiltonian system for the problem (7), (8), with a time-independent control, is given by (10) and

$$\tilde{\sigma} := \mathfrak{h}'\left(\int_0^T \nabla u \cdot \nabla q \, \mathrm{d}t \right). \tag{20}$$

Similarly, the case of a space independent coefficient $\sigma = \sigma(t)$ will lead to

$$\tilde{\sigma} := \mathfrak{h}'\Big(\frac{1}{|\Omega|}\int_\Omega \nabla u \cdot \nabla q \ \mathrm{d}x\Big),$$

and for the case where σ is constant

$$\tilde{\sigma} := \mathfrak{h}'\Big(\frac{1}{|\Omega|}\int_0^T\int_\Omega \nabla u \cdot \nabla q \ \mathrm{d}x \ \mathrm{d}t\Big).$$

3.4. **Penalty on the Control.** If we want to reconstruct a time independent control it can be beneficial to put a penalty on σ_t, *i.e.* we want to minimize the objective functional

$$F(u,\sigma_t) := \int_0^T\int_{\partial\Omega}(u-u^*)^2 \ \mathrm{d}s \ \mathrm{d}t + \varepsilon\int_0^T\int_\Omega \sigma_t^2 \ \mathrm{d}x \ \mathrm{d}t, \tag{21}$$

under the usual constraint (7). To do this the optimal control problem has to be reformulated such that σ is a state variable and the control is defined as $z := \sigma_t(x,t)$, $z : \bar{\Omega}\times[0,T] \to [z_-, z_+]$. The optimal control problem is thus to find a control z and state variables u and σ such that $F(u,z)$ is minimized and the system

$$\begin{aligned}
u_t &= \operatorname{div}(\sigma\nabla u), && \text{in } \Omega\times(0,T],\\
\sigma_t &= z && \text{in } \Omega\times(0,T],\\
\sigma\nabla u\cdot\mathbf{n} &= j, && \text{on } \partial\Omega\times(0,T],\\
u &= 0, && \text{on } \bar{\Omega}\times\{t=0\},\\
\sigma &= \sigma_0 > 0, && \text{on } \bar{\Omega}\times\{t=0\}.
\end{aligned}$$

is satisfied. The Hamiltonian becomes

$$\begin{aligned}
H(u,q,\sigma,\lambda,t) &:= \min_{z:\Omega\to[z_-,z_+]}\int_{\partial\Omega}(u-u^*)^2 \ \mathrm{d}s + \int_\Omega \operatorname{div}(\sigma\nabla u)q + z\lambda + \varepsilon z^2 \ \mathrm{d}x\\
&= \int_{\partial\Omega}(u-u^*)^2 + jq \ \mathrm{d}s - \int_\Omega \sigma\nabla u\cdot\nabla q \ \mathrm{d}x\\
&\quad + \int_\Omega \underbrace{\min_{z:\Omega\to[z_-,z_+]}\{z(\varepsilon z+\lambda)\}}_{\mathfrak{h}(\lambda)} \ \mathrm{d}x,
\end{aligned}$$

and the corresponding Hamiltonian system is

$$\begin{aligned}
u_t &= \operatorname{div}\big(\sigma\nabla u\big), && \text{in } \Omega\times(0,T],\\
\sigma_t &= \mathfrak{h}'(\lambda) && \text{in } \Omega\times(0,T],\\
\sigma\nabla u\cdot\mathbf{n} &= j, && \text{on } \partial\Omega\times(0,T],\\
u &= 0, && \text{on } \bar{\Omega}\times\{t=0\},\\
\sigma &= \sigma_0 > 0, && \text{on } \bar{\Omega}\times\{t=0\}.\\
-q_t &= \operatorname{div}\big(\sigma\nabla q\big), && \text{in } \Omega\times(0,T],\\
-\lambda_t &= -\nabla u\cdot\nabla q && \text{in } \Omega\times(0,T],\\
\sigma\nabla q\cdot\mathbf{n} &= 2(u-u^*), && \text{on } \partial\Omega\times(0,T],\\
q &= 0, && \text{on } \Omega\times\{t=T\},\\
\lambda &= 0, && \text{on } \Omega\times\{t=T\},
\end{aligned}$$

which is equivalent to (10) with

$$\tilde{\sigma} := \sigma_0 + \int_0^t \mathfrak{h}'\Big(\int_y^T -(\nabla u\cdot\nabla q)(x,z) \ \mathrm{d}z\Big) \ \mathrm{d}y.$$

Note, since we no longer have a constraint $\sigma > 0$, the bound z_- has to be carefully chosen to ensure well-posedness of the forward problem.

In a similar fashion as for penalizing temporal variations of the control it is also possible to penalize spacial variations, as was briefly mentioned in Section 3.2, where the objective was to minimize $F(u, |\nabla\sigma|)$ under the constraint (7), which leads to the Hamiltonian (15). To be able to explicitly find the minimum in the Hamiltonian we once again let σ act as a state variable, introduce the control z and the dynamics

$$\begin{aligned} \sigma_t &= \frac{z - |\nabla\sigma|^2}{\gamma}, && \text{in } \Omega \times (0, T], \\ \sigma &= \sigma_0 > 0, && \text{in } \Omega \times \{t = 0\}, \end{aligned} \tag{22}$$

for $\gamma > 0$. The slightly perturbed control problem is now to minimize the objective function $F(u, z)$ such that (7) and (22) holds, which leads to the Hamiltonian

$$\begin{aligned} H(u, q, \sigma, \lambda, t) &:= \min_{z:\Omega\to[z_-,z_+]} \int_{\partial\Omega} (u - u^*)^2 \, \mathrm{d}s + \int_\Omega \operatorname{div}(\sigma\nabla u)q + \lambda\frac{z - |\nabla\sigma|^2}{\gamma} + \varepsilon z \, \mathrm{d}x \\ &= \int_{\partial\Omega} (u - u^*)^2 + jq \, \mathrm{d}s - \int_\Omega \sigma\nabla u \cdot \nabla q + \lambda\frac{|\nabla\sigma|^2}{\gamma} \, \mathrm{d}x \\ &\quad + \int_\Omega \underbrace{\min_{z:\Omega\to[z_-,z_+]} \{z(\varepsilon + \frac{\lambda}{\gamma})\}}_{\mathfrak{h}(\lambda)} \, \mathrm{d}x, \end{aligned}$$

and the Hamiltonian system

$$\begin{aligned} u_t &= \operatorname{div}\big(\sigma\nabla u\big), \\ \sigma_t &= \mathfrak{h}'(\lambda) - \frac{|\nabla\sigma|^2}{\gamma}, \\ -q_t &= \operatorname{div}\big(\sigma\nabla q\big), \\ -\lambda_t &= \nabla u \cdot \nabla q - 2\lambda\frac{\Delta\sigma}{\gamma}. \end{aligned}$$

3.5. **Numerical Approximation and Symplectic Methods.** Let $\bar{V} \subset V := H^1(\Omega)$ be the finite element subspace of piecewise linear functions defined on a triangulation of Ω, which implies that our optimal control problems in the previous sections are approximated by optimal control problems for ordinary differential equations. We also let the functions $\mathfrak{h}_\delta$,H^δ and h_δ denote the regularized counterparts to $\mathfrak{h}$,H and h. The regularized version of $\mathfrak{h}$ is given by (12) from which the definition of H^δ follows. The regularized function h_δ can be derived from H^δ by $h_\delta := H^\delta - \langle\lambda, H^\delta_\lambda\rangle$ and a regularized version of f can be defined as $f_\delta := H^\delta_\lambda$.

Now, introduce the uniform partition $\{t_i = ki\}_{i=0}^N$, $k = T/N$ of the time interval $[0, T]$, and the corresponding finite element approximations at each time step $\varphi_n := \varphi(t_n), \lambda_n := \lambda(t_n)$. Also define a discrete regularized version $\bar{U} : \bar{V} \times [0, T] \to \mathbb{R}$ of the value function (2),

$$\bar{U}(\phi, t_m) := \min_{\varphi_m = \phi} \left\{ g(\varphi_N) + k \sum_{n=m}^{N-1} h_\delta(\varphi_n, \lambda_{n+1}) \right\},$$

where φ_n and λ_n satisfy a symplectic scheme, *e.g.* the symplectic forward Euler method

$$\begin{aligned} \varphi_{n+1} - \varphi_n &= kH^\delta_\lambda(\varphi_n, \lambda_{n+1}), && \text{for } n = m, \dots, N-1 \text{ given } \varphi_m = \phi, \\ \lambda_n - \lambda_{n+1} &= kH^\delta_\varphi(\varphi_n, \lambda_{n+1}), && \text{for } n = m, \dots, N-1 \text{ given } \lambda_N = g_\varphi(\varphi_N). \end{aligned} \tag{23}$$

Symplecticity here means that $\bar{U}_\varphi(\varphi_n, t_n) = \lambda_n$, *i.e.* the gradient of the discrete value function coincides with the discrete dual λ_n, and given that $|H - H^\delta| = \mathcal{O}(\delta)$ it can be shown that for symplectic one-step schemes

$$\left| U(\varphi_0, t_0) - g(\varphi_N) - k \sum_{n=m}^{N-1} h_\delta(\varphi_n, \lambda_{n+1}) \right| = \mathcal{O}(k),$$

for $\delta \sim k$, see [15]. It is thus essential to use a symplectic time discretization of the regularized Hamiltonian system

$$\begin{aligned} \varphi_t &= H^\delta_\lambda(\varphi, \lambda), \\ \lambda_t &= -H^\delta_\varphi(\varphi, \lambda), \end{aligned}$$

in order to have convergence in the value function.

Some examples of other symplectic schemes are the the backward Euler method

$$\begin{aligned} \varphi_{n+1} - \varphi_n &= k H^\delta_\lambda(\varphi_{n+1}, \lambda_n), \quad \text{for } n = 0, \dots, N-1 \text{ given } \varphi_0, \\ \lambda_n - \lambda_{n+1} &= k H^\delta_\varphi(\varphi_{n+1}, \lambda_n), \quad \text{for } n = 0, \dots, N-1 \text{ given } \lambda_N, \end{aligned} \tag{24}$$

and the implicit midpoint method

$$\begin{aligned} \varphi_{n+1} - \varphi_n &= k H^\delta_\lambda\left(\frac{\varphi_n + \varphi_{n+1}}{2}, \frac{\lambda_n + \lambda_{n+1}}{2} \right), \quad n = 0, \dots, N-1, \text{ given } \varphi_0, \\ \lambda_n - \lambda_{n+1} &= k H^\delta_\varphi\left(\frac{\varphi_n + \varphi_{n+1}}{2}, \frac{\lambda_n + \lambda_{n+1}}{2} \right), \quad n = 0, \dots, N-1, \text{ given } \lambda_N. \end{aligned} \tag{25}$$

See [12] for a thorough description of symplectic methods.

3.6. **The Newton Method.** To solve the coupled nonlinear symplectic schemes (23)-(25) above, it is tempting to propose fix-point schemes that partly removes the coupling between the forward and bacward equation, *e.g.* by iterating separately in φ and λ. Such methods has the advantage that existing partial differential equation solvers can be used to efficiently solve the forward and backward problems in each iteration, but the disadvantage is that the convergence to an optimal solution tends to be slow, and also dependent on the discretization. A more suitable strategy is to use information of the Hessian of H^δ; *e.g.* Quasi-Newton methods, or since the Hessian in our case can be found explicitly and is sparse, the Newton method itself.

For the Hamiltonian system (10) with $\tilde{\sigma} := \mathfrak{h}'_\delta$ given by (12) the symplectic backward Euler can be written as

$$F_n(w) = 0, \quad G_n(w) = 0, \quad n = 0, \dots, N-1, \quad \forall w \in \bar{V}$$

where

$$\begin{aligned} F_n(w) &:= \int_\Omega (u_{n+1} - u_n) w + k \mathfrak{h}'_\delta(\nabla u_{n+1} \cdot \nabla q_n) \nabla u_{n+1} \cdot \nabla w \, \mathrm{d}x \\ &\quad - \int_{\partial\Omega} k j_{n+1} w \, \mathrm{d}s, \\ G_n(w) &:= \int_\Omega (q_n - q_{n+1}) w + k \mathfrak{h}'_\delta(\nabla u_{n+1} \cdot \nabla q_n) \nabla q_n \cdot \nabla w \, \mathrm{d}x \\ &\quad - \int_{\partial\Omega} 2k(u_{n+1} - u^*_{n+1}) w \, \mathrm{d}s, \end{aligned} \tag{26}$$

and $u_0 = q_N = 0$. Given an initial guess $u[0]$, $q[0]$ the (damped) Newton method yields that

$$\begin{aligned} u[i+1] &= u[i] - \alpha \hat{u}, \\ q[i+1] &= q[i] - \alpha \hat{q}, \end{aligned}$$

where $\alpha \in (0, 1]$ and, for each iteration, the updates $\hat{u}$ and $\hat{q}$ solve a linear system of the form

$$\begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{11}^T \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{q} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{27}$$

where

$$\hat{u} = \begin{pmatrix} \hat{u}_1 & \dots & \hat{u}_N \end{pmatrix}^T, \qquad \hat{q} = \begin{pmatrix} \hat{q}_0 & \dots & \hat{q}_{N-1} \end{pmatrix}^T,$$
$$f = \begin{pmatrix} F_0 & \dots & F_{N-1} \end{pmatrix}^T, \quad g = \begin{pmatrix} G_0 & \dots & G_{N-1} \end{pmatrix}^T.$$

The matrix K_{11} is a bi-diagonal block matrix with $M + S_i$ for $i = 0, \dots, N-1$ on the diagonal and $-M$ on the sub-diagonal, where M denotes the mass matrix

$$\int_\Omega w\bar{w} \, \mathrm{d}x,$$

and

$$S_n := \int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+1} \cdot \nabla q_n)\nabla q_n \cdot \nabla w \ \nabla u_{n+1} \cdot \nabla \bar{w} \, \mathrm{d}x + \int_\Omega k\mathfrak{h}_\delta'(\nabla u_{n+1} \cdot \nabla q_n)\nabla w \cdot \nabla \bar{w} \, \mathrm{d}x.$$

for $w, \bar{w} \in \bar{V}$. The matrices K_{12}, K_{21} are symmetric block-diagonal matrices with

$$\int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+1} \cdot \nabla q_n)\nabla u_{n+1} \cdot \nabla w \ \nabla u_{n+1} \cdot \nabla \bar{w} \, \mathrm{d}x,$$

and

$$\int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+1} \cdot \nabla q_n)\nabla q_n \cdot \nabla w \ \nabla q_n \cdot \nabla \bar{w} \, \mathrm{d}x - \int_{\partial\Omega} 2k\bar{w}w \, \mathrm{d}s,$$

for $n = 0, \dots, N-1$ on the the diagonals, respectively.

If we repartition the block 2×2 linear system (27) to

$$\begin{pmatrix} K_{21} & K_{11}^T \\ K_{11} & K_{12} \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{q} \end{pmatrix} = \begin{pmatrix} g \\ f \end{pmatrix}, \tag{28}$$

we see that it is a generalized saddle point system [4] with symmetric matrices K_{21}, K_{12}, and $K_{11}^T \neq 0$, $K_{21} \neq 0$. However, unlike saddle point problems arising from *e.g.* the steady-state Navier-Stokes equations or from the Karush-Kuhn-Tucker optimality conditions for equality constrained minimization problems, both K_{12} and K_{21} may here be indefinite and singular.

Since (27) and (28) are increasingly ill-conditioned with respect to reduction in mesh size, step size and regularization, the success of iterative algorithms like Krylov sub-space methods will depend heavily on the choice of preconditioner. Standard algebraic preconditioners like incomplete LU-factorization are often unsuitable for saddle-point problems due to the indefiniteness and lack of diagonal dominance, so the preconditioner must be tailored for the specific problem at hand. One popular approach for PDE-constrained optimization problems is to base the preconditioner on the solution from a reduced approximated problem where the Schur complement is replaced by an approximation *e.g.* by quasi-newton methods, see [5].

In our case we use the GMRES method to solve the non-symmetric system (27) and base our preconditioner on the approximate solution of a simple blockwise Gauss-Seidel method *i.e.* to start with a guess $\hat{q}^0$ and iteratively solve

$$\begin{aligned} K_{11}\hat{u}^{i+1} &= f - K_{12}\hat{q}^i, \\ K_{11}^T\hat{q}^{i+1} &= g - K_{21}\hat{u}^{i+1}, \end{aligned} \tag{29}$$

which works well for large regularizations *i.e.* when $\mathfrak{h}_\delta''$ is small and the diagonal blocks of (27) are dominant. Also, each iteration with this method only requires

one forward and one backward solve in time of a modified heat equation so the computational work for one iteration is concentrated to solving $N-1$ smaller systems with system matrices $(M+S_i)$. In practice, the Gauss-Seidel method will break down for small regularizations but for our problems (and discretizations) only one iteration with (29) turns out to be a fairly good approximation to use as preconditioner. Note that for $\hat{q}^0 = 0$, one Gauss-Seidel iteration is the same as solving (27) with the approximation $K_{12} = 0$.

Another more elaborate idea is to use a preconditioner based on the solution of an approximated Schur complement system

$$\begin{pmatrix} K_{11} & K_{12} \\ 0 & S \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{q} \end{pmatrix} = \begin{pmatrix} g \\ f - K_{12}K_{11}^{-1}g \end{pmatrix},$$

where S is an approximation of the Schur complement

$$K_{11}^T - K_{21}K_{11}^{-1}K_{12}.$$

which essentially is to find a good approximation of the lower triangular block matrix K_{11}^{-1}.

Although solution algorithms for saddle point systems on the symmetric form (28) are extensively treated in the litterature, see [4] for an overview, we here favour the non-symmetric form (27), since a Schur complement reduction of (28) means to find an approximation to the Schur complement

$$K_{12} - K_{11}K_{21}^{-1}K_{11}^T,$$

which since K_{21} here can be singular, is unavailable. One way around this obstacle is to rewrite (28) by *e.g.* the augmented Lagrangian method which leads to a symmetric invertible Schur complement but where the physical meaning of the original system, on PDE level, is partially lost.

If a direct solver is used for the Newton system it is appropriate to reorder (27) such that the solution vector and right hand side contains time steps in increasing order, which leads to a banded Jacobian with band-width of the same order as the number of spacial degrees of freedom.

Our computations were implemented MATLAB (for the one dimensional examples), and in DOLFIN [13], the C++/Python interface of the finite element solver environment FEniCS [11] (for the two dimensional examples). Piecewise linear basis functions were used for the finite element subspace $\bar{V}$, and in all examples the solution u, q was first calculated for a large regularization which was succesively reduced such that the solution from the previous regularization served as starting guess for a smaller regularization.

For the two dimensional examples the sadde-point system (27) was solved with the PETSc implementation of GMRES (used by DOLFIN) with preconditioning based on the solution from one iteration of blockwise Gauss-Seidel method. For the one dimensional examples a direct solver was used. The number of iterations for GMRES with the Gauss-Seidel preconditioner seems to be relatively insensitive with respect to temporal and spacial discretization but still highly sensitive to the regularization in our examples.

To give a time independent approximation $\sigma(x)$ of the time dependent control $\sigma(x,t)$, approximated by $\tilde{\sigma} := \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)$ where u, q are solutions to the Hamiltonian system (10), three different types of averaging were tested as post-processing:

(1) Let the time independent control be defined by the Hamiltonian (18), *i.e.*

$$\sigma := \mathfrak{h}'_\delta\Big(\int_0^T \nabla u \cdot \nabla q \, \mathrm{d}t \Big). \tag{30}$$

(2) Let the time independent control be the average of the time dependent control, *i.e.*

$$\sigma := \frac{1}{T}\int_0^T \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)\ \mathrm{d}t. \tag{31}$$

(3) Let the time independent control be the weighted average

$$\sigma := \frac{\int_0^T \mathfrak{h}'_\delta(\nabla u \cdot \nabla q)|\nabla u \cdot \nabla q|\ \mathrm{d}t}{\int_0^T |\nabla u \cdot \nabla q|\ \mathrm{d}t}, \tag{32}$$

of the time dependent control $\mathfrak{h}'_\delta(\nabla u \cdot \nabla q)$.

The weighted average turned out to be the most successful aproximation and can be explained by first extending the Hamiltonian (9) to also depend on the artifical variable z as in Section 3.3

$$\begin{aligned} H(u,q,z) :=& \frac{1}{\bar{T}}\int_0^T \int_{\partial\Omega} (u-u^*)^2 + jq\ \mathrm{d}s\ \mathrm{d}t - \frac{1}{\bar{T}}\int_0^T \int_\Omega u_t q\ \mathrm{d}x\ \mathrm{d}t \\ & - \frac{1}{\bar{T}}\int_\Omega \int_0^T \mathfrak{h}'(\nabla u \cdot \nabla q)\nabla u \cdot \nabla q\ \mathrm{d}t\ \mathrm{d}x, \end{aligned}$$

where $\mathfrak{h}'(\nabla u \cdot \nabla q)\nabla u \cdot \nabla q = \mathfrak{h}(\nabla u \cdot \nabla q)$ by definition. For the problem with a time independent control we now seek an approximation of the Hamiltonian (18) of the form

$$\begin{aligned} \bar{H}(u,q,z) :=& \frac{1}{\bar{T}}\int_0^T \int_{\partial\Omega} (u-u^*)^2 + jq\ \mathrm{d}s\ \mathrm{d}t - \frac{1}{\bar{T}}\int_0^T \int_\Omega u_t q\ \mathrm{d}x\ \mathrm{d}t \\ & - \frac{1}{\bar{T}}\int_\Omega f(\nabla u \cdot \nabla q)\int_0^T \nabla u \cdot \nabla q\ \mathrm{d}t\ \mathrm{d}x, \end{aligned}$$

that best approximates H, *i.e.*

$$f(\nabla u \cdot \nabla q) := \frac{\int_0^T \mathfrak{h}'(\nabla u \cdot \nabla q)\nabla u \cdot \nabla q\ \mathrm{d}t}{\int_0^T \nabla u \cdot \nabla q\ \mathrm{d}t}.$$

In Figure 3, one dimensional reconstructions from three sets of simulated data u^*, generated from a time independent conductivity σ_{true}, are compared:

(1) Data calculated with the same discretization as u and q.
(2) Different discretizations used for data and solutions.
(3) Different discretizations used for data and solutions and with noise in the data u^*.

The last set is the most realistic one since for true experimental data of u^* it is inevitable to have noisy measurements. To simulate noise the discrete solution u^* was multiplied componentwise by independent standard normal distributed stochastic variables ε_{ij} according to $u^*(x_i,t_j)(1+\eta\epsilon_{ij})$, where η denotes the percentage of noise. It is notable that the systematic error from using different meshes can have a much bigger effect on the solutions than additional noise, which can be observed from the dual solution q in Figure 3.

In Figure 4 the time independent post-processing of the time dependent reconstruction can be found. It is here evident that the weighted average (32) performs better than (31), but since the reconstruction is highly dependent on the given boundary condition, see Figure 5 for comparison, there are situations where the different post-processing techniques perform equally well. It would of course be optimal to use the knowledge that σ_{true} is independent of time in the calculations, *i.e.* to use the Newtion system for (10) with time independent-control (20). This would however lead to a dense Jacobian.

Note that in the examples the limits σ_- and σ_+ were chosen to be the biggest and smallest values of σ_{true}. In our experience the Pontryagin method is not well suited for reconstruction of values between σ_- and σ_+, if there is noise or other measurement errors present in data.

Figure 6 shows two-dimensional reconstructions of two different time independent conductivities. Unlike the one-dimensional example the quality of the reconstruction here deteriorates quickly as the distance to the measurement locations is increased.

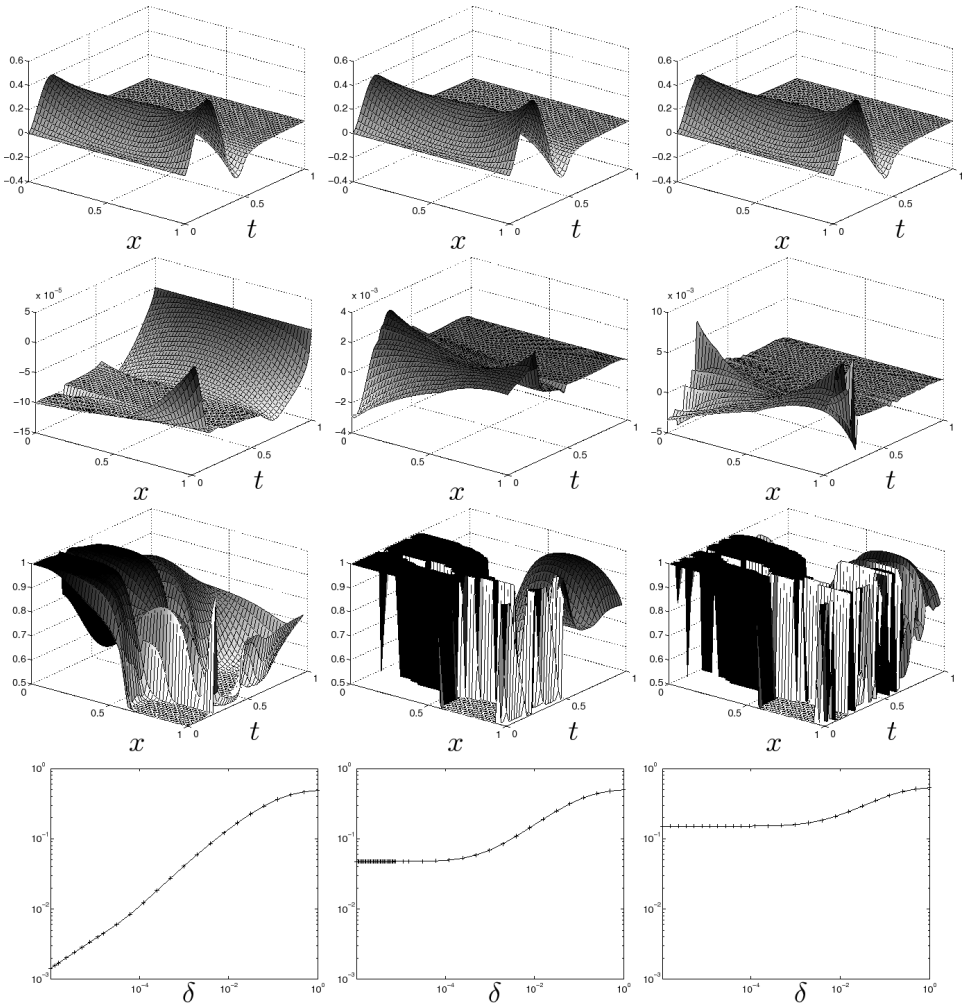

FIGURE 3. 1D reconstruction of $\sigma_{true} = 0.75 - 0.5\tanh(20x - 10)$ for $\delta = 10^{-6}$, $\sigma_- = 0.5$, and $\sigma_+ = 1$. Measurements were collected on both boundaries and the Neumann boundary condition was $\sigma u_x(0,t) = -\sigma u_x(1,t) = \sin(4t)$ for $t < 0.5$ and 0 elsewhere. The plot shows, from top to bottom, u, q, $\mathfrak{h}'_\delta$ and the objective function $\|u - u^*\|_{L^2(\partial\Omega\times[0,T])}$. In all cases u, q was calculated with 50 steps in space and time. In the left column, the data u^* was generated by solving the heat equation for σ_{true} with 50 steps in time and space, while 200 steps in time and space was used in the middle and right columns. In the right column 10% noise was also added to u^*.

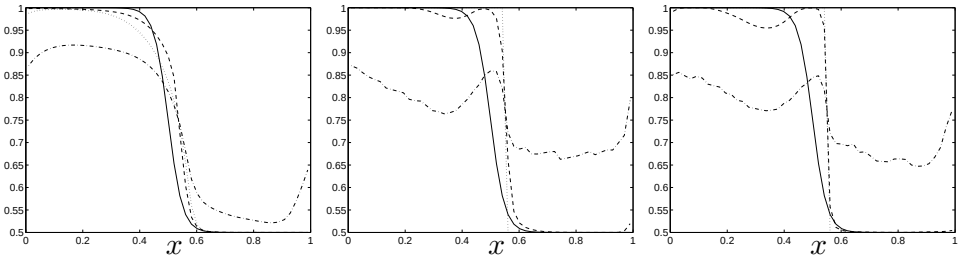

FIGURE 4. The time independent post-processed conductivity for the 1D reconstructions in Figure 3. The true control σ_{true} is indicated by a solid line and the averaged controls (30), (31) and (32) are indicated by dotted, dash-dotted and dashed lines, respectively.

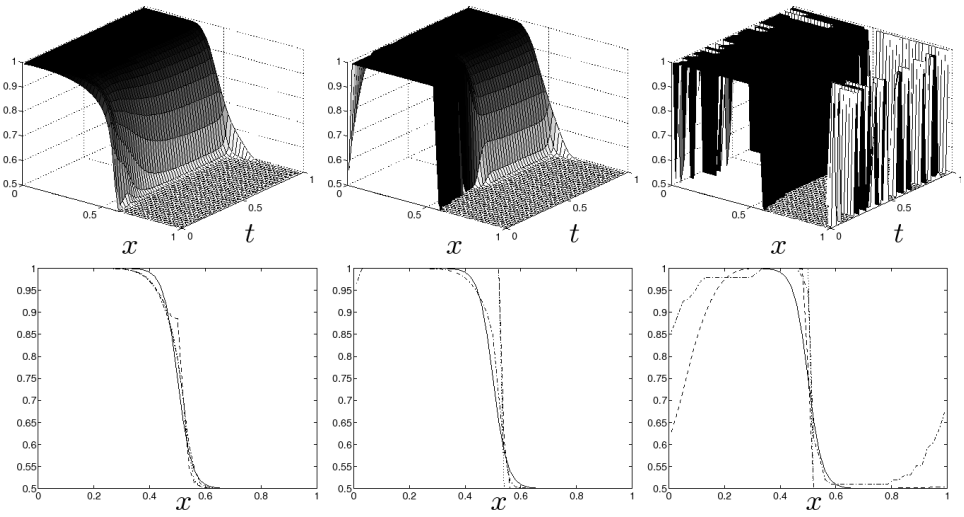

FIGURE 5. 1D reconstruction with data as in Figure 3 and 4 but with Neumann boundary condition $\sigma u_x(0,t) = -\sigma u_x(1,t) = 1$. The top row shows $\mathfrak{h}'_\delta$ and the bottom row the averaged conductivities, as described in Figure 4.

4. Reconstruction from the Wave Equation

In this section the goal is to determine the wave speed for a scalar acoustic wave equation: Given measured data u^*, find the state $u : \bar{\Omega} \times [0,T] \to \mathbb{R}$, $u(\cdot,t) \in V$ and a control $\sigma : \bar{\Omega} \times [0,T] \to [\sigma_-, \sigma_+]$, $\sigma = \sigma(x,t)$ where $0 < \sigma_- < \sigma_+$, that solves the partial differential equation

$$
\begin{aligned}
u_{tt} &= \operatorname{div}(\sigma \nabla u), && \text{in } \Omega \times (0,T], \\
\sigma \nabla u \cdot \mathbf{n} &= j, && \text{on } \partial\Omega \times (0,T], \\
u &= u_t = 0, && \text{on } \bar{\Omega} \times \{t = 0\},
\end{aligned}
\tag{33}
$$

such that the error functional

$$
\int_0^T \int_{\partial\Omega} (u - u^*)^2 \, \mathrm{d}s \, \mathrm{d}t, \tag{34}
$$

is minimized. The control σ is here the square of the wave speed of the medium and u is the pressure deviation.

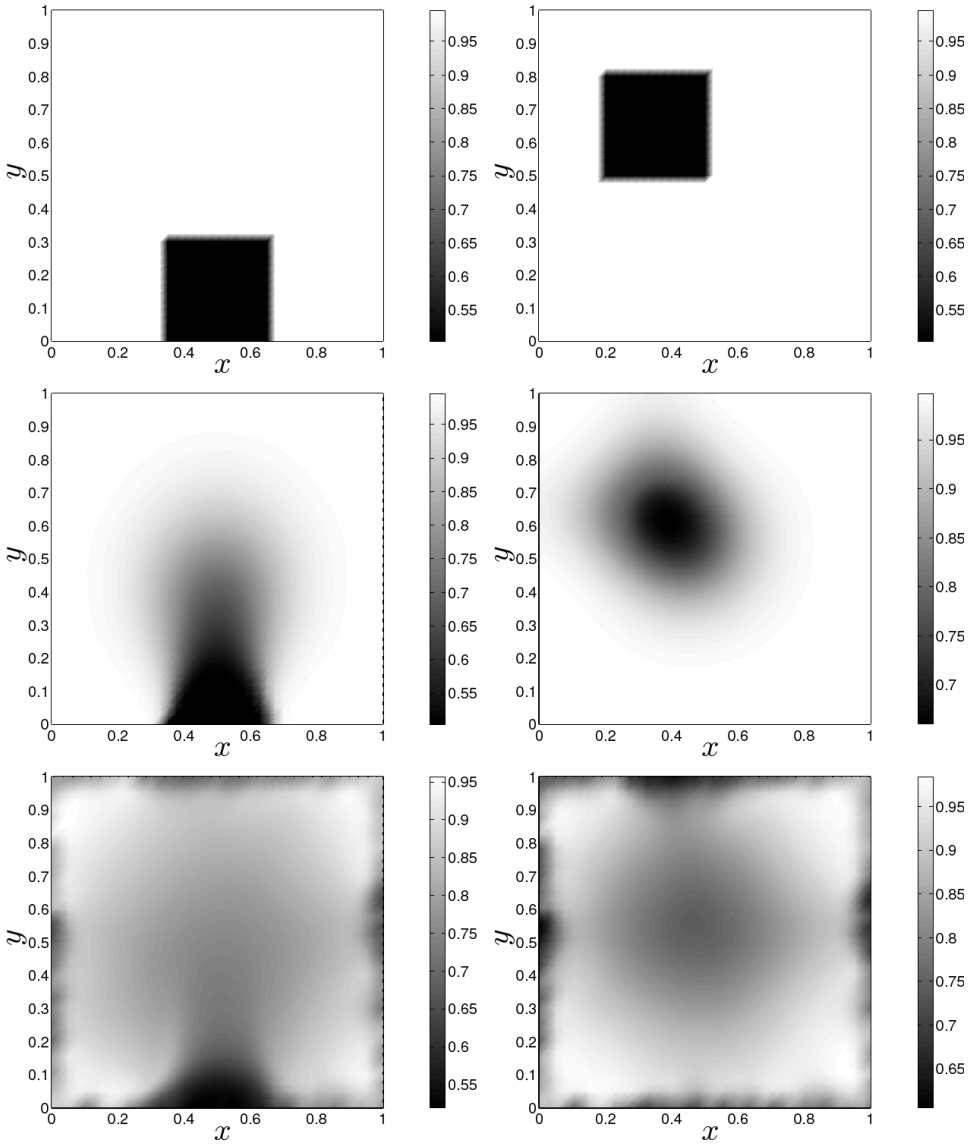

FIGURE 6. 2D reconstruction on the unit square with final time $T = 1$ and Neumann boundary condition $\sigma \frac{\partial u}{\partial n} = 1$ on $\partial\Omega \times [0, T]$. The data u^* was simulated by solving the forward equation on a quasi-uniform mesh with 13000 triangles and 80 time steps while the inverse problem was solved on a uniform mesh with 3200 triangles and 40 time steps. Measurements from the whole boundary were used. Top: True conductivity σ_{true}. Middle: Reconstructed condictivity for $\delta \approx 0.002$ using the weighted average (32). Bottom: As in middle but for $\delta \approx 0.05$ and with 5% noise in the measurements.

To use the framework of the previous section we note that (33) can be written as the first order system

$$
\begin{aligned}
v_t &= \operatorname{div}(\sigma \nabla u), && \text{in } \Omega \times (0, T], \\
u_t &= v, && \text{in } \bar{\Omega} \times (0, T], \\
\sigma \nabla u \cdot \mathbf{n} &= j, && \text{on } \partial\Omega \times (0, T], \\
u &= v = 0, && \text{on } \bar{\Omega} \times \{t = 0\}.
\end{aligned} \tag{35}
$$

4.1. **The Hamiltonian System.** As in Section 3.1 we have a Hamiltonian associated with the optimal control problem (34) and (35) which is defined by

$$\begin{aligned} H &:= \min_{\sigma:\Omega\to[\sigma_-,\sigma_+]} \int_{\partial\Omega} (u-u^*)^2 \,\mathrm{d}s + \int_\Omega \operatorname{div}(\sigma\nabla u)q + vp \,\mathrm{d}x \\ &= \int_{\partial\Omega} (u-u^*)^2 + jq \,\mathrm{d}s + \min_{\sigma:\Omega\to[\sigma_-,\sigma_+]} \int_\Omega -\sigma\nabla u\cdot\nabla q + vp \,\mathrm{d}x \\ &= \int_{\partial\Omega} (u-u^*)^2 + jq \,\mathrm{d}s + \int_\Omega vp - \underbrace{\max_{\sigma\in[\sigma_-,\sigma_+]} \{\sigma\nabla u\cdot\nabla q\}}_{\mathfrak{h}(\nabla u\cdot\nabla q)} \,\mathrm{d}x, \end{aligned} \tag{36}$$

and the Hamiltonian system becomes

$$\begin{aligned} v_t &= \operatorname{div}\big(\tilde\sigma\nabla u\big), && \text{in } \Omega\times(0,T], \\ u_t &= v, && \text{in } \bar\Omega\times(0,T], \\ \tilde\sigma\nabla u\cdot\mathbf{n} &= j, && \text{on } \partial\Omega\times(0,T], \\ u &= v = 0, && \text{on } \bar\Omega\times\{t=0\}, \\ -p_t &= \operatorname{div}\big(\tilde\sigma\nabla q\big), && \text{in } \Omega\times(0,T], \\ -q_t &= p, && \text{in } \bar\Omega\times(0,T], \\ \tilde\sigma\nabla q\cdot\mathbf{n} &= 2(u-u^*), && \text{on } \partial\Omega\times(0,T], \\ p &= q = 0, && \text{on } \Omega\times\{t=T\}, \end{aligned} \tag{37}$$

or equivalently

$$\begin{aligned} u_{tt} &= \operatorname{div}\big(\tilde\sigma\nabla u\big), && \text{in } \Omega\times(0,T], \\ \tilde\sigma\nabla u\cdot\mathbf{n} &= j, && \text{on } \partial\Omega\times(0,T], \\ u &= u_t = 0, && \text{on } \bar\Omega\times\{t=0\}, \\ q_{tt} &= \operatorname{div}\big(\tilde\sigma\nabla q\big), && \text{in } \Omega\times(0,T], \\ \tilde\sigma\nabla q\cdot\mathbf{n} &= 2(u-u^*), && \text{on } \partial\Omega\times(0,T], \\ q &= q_t = 0, && \text{on } \bar\Omega\times\{t=T\}. \end{aligned} \tag{38}$$

with

$$\tilde\sigma := \mathfrak{h}'(\nabla u\cdot\nabla q).$$

4.2. **Symplecticity for the Wave Equation.** As a natural case the symplectic methods discussed in 3.5, with $\varphi=(u,v)$, $\lambda=(p,q)$, can be used to solve the system (37). It is however also possible to use a time-discretization that is symmetric in time i.e.

$$\begin{aligned} u_{n+1}-2u_n+u_{n-1} &= k^2\operatorname{div}\big(\tilde\sigma_n\nabla u_n\big), && \text{in } \Omega, \\ \tilde\sigma_n\nabla u_n\cdot\mathbf{n} &= j_n, && \text{on } \partial\Omega, \\ u_0 = u_1 &= 0, && \text{in } \Omega, \\ q_{n+1}-2q_n+q_{n-1} &= k\operatorname{div}\big(\tilde\sigma_n\nabla q_n\big), && \text{in } \Omega, \\ \tilde\sigma_n\nabla q_n\cdot\mathbf{n} &= 2(u_n-u_n^*), && \text{on } \partial\Omega, \\ q_N = q_{N-1} &= 0, && \text{in } \Omega, \end{aligned} \tag{39}$$

for $\tilde\sigma_n := \mathfrak{h}'(\nabla u_n\cdot\nabla q_n)$ and $n=1,\dots,N-1$. For a given $\tilde\sigma$, constant in time, this scheme is the symplectic backward Euler method for the forward wave equation for u, which can be written as the Hamiltonian system (35) with Hamiltonian

$$H_{wave}(u,v) := \frac{1}{2}\int_\Omega |\tilde\sigma\nabla u|^2 + v^2 \,\mathrm{d}x,$$

and the symplectic forward Euler method for the backward wave equation for q.

To see that that the scheme (39) is symplectic for $\tilde{\sigma}_n := \mathfrak{h}'(\nabla u_n \cdot \nabla q_n)$ we note that a one-step method $(\varphi_n, \lambda_n) \to (\varphi_{n+1}, \lambda_{n+1})$ is symplectic if there exists a function $H(\varphi_n, \lambda_{n+1})$ such that (23) holds, or equivalently $H(\varphi_{n+1}, \lambda_n)$ such that (24) holds, see Remark 4.8 in [15] or [12] for details. It thus follows that the one-step method

$$\begin{aligned} v_{n+1} - v_n &= k\mathrm{div}\big(\mathfrak{h}'(\nabla u_n \cdot \nabla q_n)\nabla u_n\big), \\ u_{n+1} - u_n &= k v_{n+1}, \\ p_n - p_{n+1} &= k\mathrm{div}\big(\mathfrak{h}'(\nabla u_n \cdot \nabla q_n)\nabla q_n\big), \\ q_n - q_{n+1} &= k p_{n+1}, \end{aligned}$$

corresponds to the symplectic forward Euler method for the Hamiltonian

$$\tilde{H}(\underbrace{u_n, q_n}_{\varphi_n}, \underbrace{v_{n+1}, p_{n+1}}_{\lambda_{n+1}}) := H(u_n, v_{n+1}, p_{n+1}, q_n) - 2\int_\Omega v_{n+1}p_{n+1}\ \mathrm{d}x,$$

where H is given by (36). Since (39) only is stable for sufficiently small time-steps and still requires to solve a complex saddle point system we will use the symplectic midpoint method in our experiments.

4.3. **Numerical Examples.** Let $\tilde{\sigma} := \mathfrak{h}'_\delta$ where $\mathfrak{h}'_\delta$ is given by (12). The symplectic midpoint method for the regularized Hamiltonian system (37) can then be written as

$$F_n^1(w) = 0, \quad F_n^2(w) = 0, \quad G_n^1(w) = 0, \quad G_n^2(w) = 0,$$

for $n = 0, \dots, N-1$, and $\forall w \in \bar{V}$, where

$$\begin{aligned} F_n^1(w) &:= \int_\Omega (v_{n+1} - v_n)w + k\mathfrak{h}'_\delta\big(\nabla u_{n+\frac{1}{2}} \cdot \nabla q_{n+\frac{1}{2}}\big)\nabla u_{n+\frac{1}{2}} \cdot \nabla w\ \mathrm{d}x \\ &\quad - \int_{\partial\Omega} k j_{n+\frac{1}{2}} w\ \mathrm{d}s, \\ F_n^2(w) &:= \int_\Omega (u_{n+1} - u_n - k v_{n+\frac{1}{2}})w\ \mathrm{d}x, \\ G_n^1(w) &:= \int_\Omega (q_n - q_{n+1} - k p_{n+\frac{1}{2}})w\ \mathrm{d}x. \\ G_n^2(w) &:= \int_\Omega (p_n - p_{n+1})w + k\mathfrak{h}'_\delta(\nabla u_{n+\frac{1}{2}} \cdot \nabla q_{n+\frac{1}{2}})\nabla q_{n+\frac{1}{2}} \cdot \nabla w\ \mathrm{d}x \\ &\quad - \int_{\partial\Omega} 2k(u_{n+\frac{1}{2}} - u^*_{n+\frac{1}{2}})w\ \mathrm{d}s, \end{aligned}$$

and $u_0 = v_0 = p_N = q_N = 0$. The index $n + \frac{1}{2}$ implies the average of the values at n and $n+1$, *i.e.* $u_{n+\frac{1}{2}} := \frac{1}{2}(u_n + u_{n+1})$. Taking the variations with respect to u, v, p, q gives the Newton system

$$\begin{pmatrix} K_{11} & K_{12} & 0 & K_{14} \\ K_{21} & K_{22} & 0 & 0 \\ 0 & 0 & K_{33} & K_{34} \\ K_{41} & 0 & K_{43} & K_{44} \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{p} \\ \hat{q} \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ g_1 \\ g_2 \end{pmatrix}, \tag{40}$$

with increments

$$\begin{aligned} \hat{u} &= \begin{pmatrix} \hat{u}_1 & \dots & \hat{u}_N \end{pmatrix}^T, \qquad \hat{v} = \begin{pmatrix} \hat{v}_1 & \dots & \hat{v}_N \end{pmatrix}^T, \\ \hat{p} &= \begin{pmatrix} \hat{p}_0 & \dots & \hat{p}_{N-1} \end{pmatrix}^T, \quad \hat{q} = \begin{pmatrix} \hat{q}_0 & \dots & \hat{q}_{N-1} \end{pmatrix}^T, \end{aligned}$$

and right hand side

$$f_1 = \begin{pmatrix} F_0^1 & \dots & F_{N-1}^1 \end{pmatrix}^T, \quad f_2 = \begin{pmatrix} F_0^2 & \dots & F_{N-1}^2 \end{pmatrix}^T,$$
$$g_1 = \begin{pmatrix} G_0^1 & \dots & G_{N-1}^1 \end{pmatrix}^T, \quad g_2 = \begin{pmatrix} G_0^2 & \dots & G_{N-1}^2 \end{pmatrix}^T.$$

with submatrices with the following structure:

- K_{11} is lower block bi-diagonal with

$$\begin{aligned} &\frac{1}{2}\int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+\frac{1}{2}}\cdot\nabla q_{n+\frac{1}{2}})\nabla q_{n+\frac{1}{2}}\cdot\nabla w\ \nabla u_{n+\frac{1}{2}}\cdot\nabla\bar{w}\ \mathrm{d}x \\ &+\frac{1}{2}\int_\Omega k\mathfrak{h}_\delta'(\nabla u_{n+\frac{1}{2}}\cdot\nabla q_{n+\frac{1}{2}})\nabla w\cdot\nabla\bar{w}\ \mathrm{d}x, \end{aligned} \tag{41}$$

  on its main diagonal for $n = 0, \dots, N-1$ and on its sub-diagonal for $n = 1, \dots, N-1$.
- K_{44} is upper block bi-diagonal with (41) on its diagonal for $n = 0, \dots, N-1$ and on its super-diagonal for $n = 0, \dots, N-2$.
- $K_{12} = K_{21} = K_{34}^T = K_{43}^T$ is lower block bi-diagonal with mass matrices M on the main diagonal and $-M$ on the subdiagonal.
- $K_{22} = K_{33}^T$ is lower block bi-diagonal with $-\frac{kM}{2}$ on the diagonal and the sub-diagonal.
- K_{14} is upper block bi-diagonal with

$$\frac{1}{2}\int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+\frac{1}{2}}\cdot\nabla q_{n+\frac{1}{2}})\nabla u_{n+\frac{1}{2}}\cdot\nabla w\ \nabla u_{n+\frac{1}{2}}\cdot\nabla\bar{w}\ \mathrm{d}x,$$

  on its diagonal for $n = 0, \dots, N-1$ and on its super-diagonal for $n = 0, \dots, N-2$.
- K_{41} is lower block bi-diagonal with

$$\frac{1}{2}\int_\Omega k\mathfrak{h}_\delta''(\nabla u_{n+\frac{1}{2}}\cdot\nabla q_{n+\frac{1}{2}})\nabla q_{n+\frac{1}{2}}\cdot\nabla w\ \nabla q_{n+\frac{1}{2}}\cdot\nabla\bar{w}\ \mathrm{d}x - \int_{\partial\Omega} k\bar{w}w\ \mathrm{d}s,$$

  on its diagonal for $n = 0, \dots, N-1$ and sub-diagonal for $n = 1, \dots, N-1$.

As in the previous section we will solve the Newton system using GMRES and an approximate solution as preconditioner, *e.g.* from the the 2×2 blockwise Gauss-Seidel method

$$\begin{aligned} K_{11}\hat{u}^{i+1} + K_{12}\hat{v}^{i+1} &= f_1 - K_{14}\hat{q}^i, \\ K_{21}\hat{u}^{i+1} + K_{22}\hat{v}^{i+1} &= f_2, \\ K_{33}\hat{p}^{i+1} + K_{34}\hat{q}^{i+1} &= g_1, \\ K_{43}\hat{p}^{i+1} + K_{44}\hat{q}^{i+1} &= g_2 - K_{41}\hat{u}^{i+1}, \end{aligned}$$

which can be written as

$$\begin{aligned} (K_{11} - K_{12}K_{22}^{-1}K_{21})\hat{u}^{i+1} &= f_1 - K_{12}K_{22}^{-1}f_2 - K_{14}\hat{q}^i, \\ (K_{44} - K_{43}K_{33}^{-1}K_{34})\hat{q}^{i+1} &= g_2 - K_{43}K_{33}^{-1}g_1 - K_{41}\hat{u}^{i+1}. \end{aligned} \tag{42}$$

Note that (42) is easily solved since inverting K_{22} and K_{33} only involves the calculation of M^{-1}. In fact, the Schur complements $K_{11} - K_{12}K_{22}^{-1}K_{21}$ and $K_{44} - K_{43}K_{33}^{-1}K_{34}$ becomes lower and upper block trianglar matrices, respectively, and (42) can be solved by one forward substitution in time for $\hat{u}^{i+1}$ and one backward substitution in time for $\hat{q}^{i+1}$. Of course, to save memory the Schur complement system (42) should never be formed explicitly. For large regularizations the Schur complements can be seen as approximations of the operator $-\Delta + \partial_{tt}$. As for the case with the heat equation starting with $\hat{q}^0 = 0$, one iteration with (42) is the same as solving (40) with $K_{14} = 0$.

In Figure 7, a two dimensional example of reconstruction two different speed coefficients is shown. The measured data was here simulated by solving the wave

equation for σ_{true} with the symplectic backward Euler method for (35), which can be written as the second order scheme

$$\int_\Omega (u_{n+1} - 2u_n + u_{n-1})w \, \mathrm{d}x = \int_{\partial\Omega} jw \, \mathrm{d}s - \int_\Omega \sigma \nabla u_n \cdot \nabla w \, \mathrm{d}x, \quad \forall w \in \bar{V}.$$

Since the wave equation is a conservation law and is reversible in time it is tempting to believe that it would be easier to control than the heat equation but there are some computational drawbacks: numerical errors are propagated in time and there seems to be many local minima. From the approximation $\mathfrak{h}'_\delta(\nabla u \cdot \nabla q)$ in Figure 8 it is evident that the time dependent reconstruction varies a lot over time and is not a good approximation of the time independent wave coefficient σ_{true}.

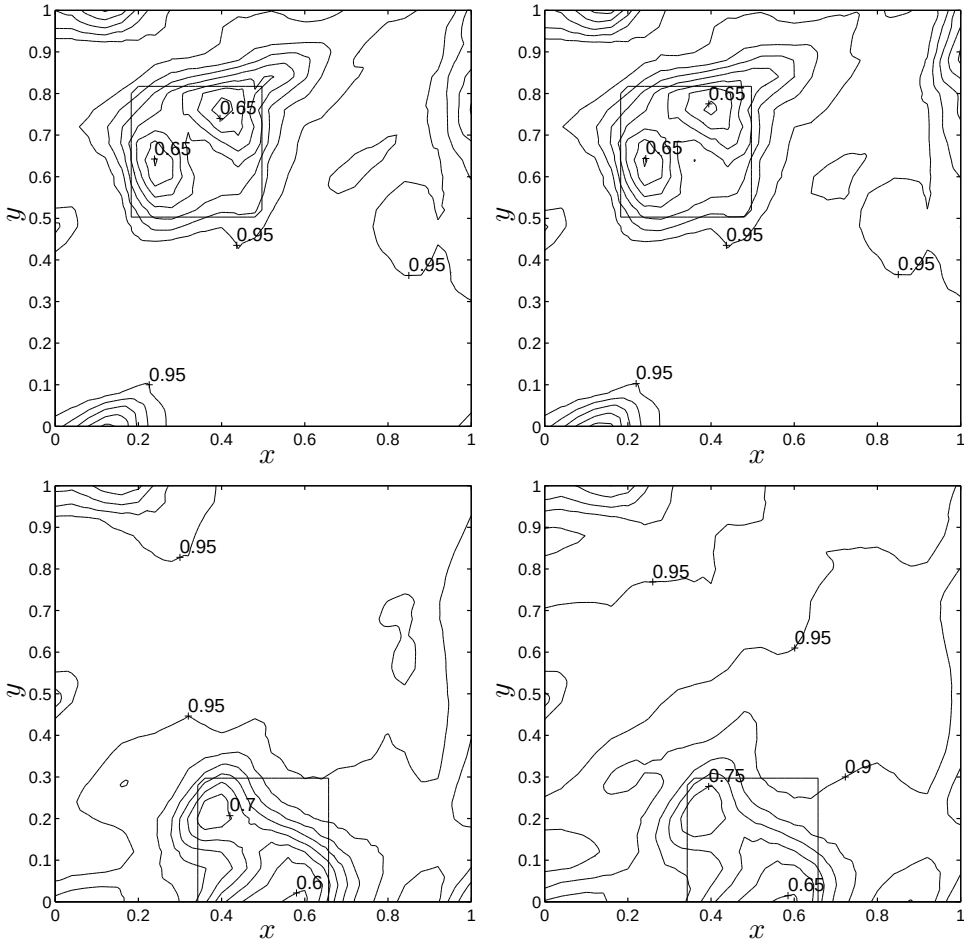

FIGURE 7. 2D reconstruction using the weighted average (32), final time $T = 1.5$ and Neumann boundary condition $2\sin(4\pi t)$ for $(x, y, t) \in \{0\} \times [0.4, 0.6] \times [0, 0.5]$ and 0 elsewhere. The data u^* was simulated by solving the forward equation on a quasi-uniform mesh with 3232 triangles and 328 time steps while the inverse problem was solved on a uniform mesh with 1250 triangles and 30 time steps. Measurements from the whole boundary were used. Top: Reconstruction of $\sigma_{true} = 0.5$ inside the square $[0.2, 0.5] \times [0.5, 0.8]$ and $\sigma_{true} = 1$ elsewhere, with no noise in data (left) and 10% noise in data (right). Bottom: Reconstruction of $\sigma_{true} = 0.5$ inside the square $[0.35, 0.65] \times [0, 0.3]$ and $\sigma_{true} = 1$ elsewhere, with no noise in data (left) and 10% noise in data (right).

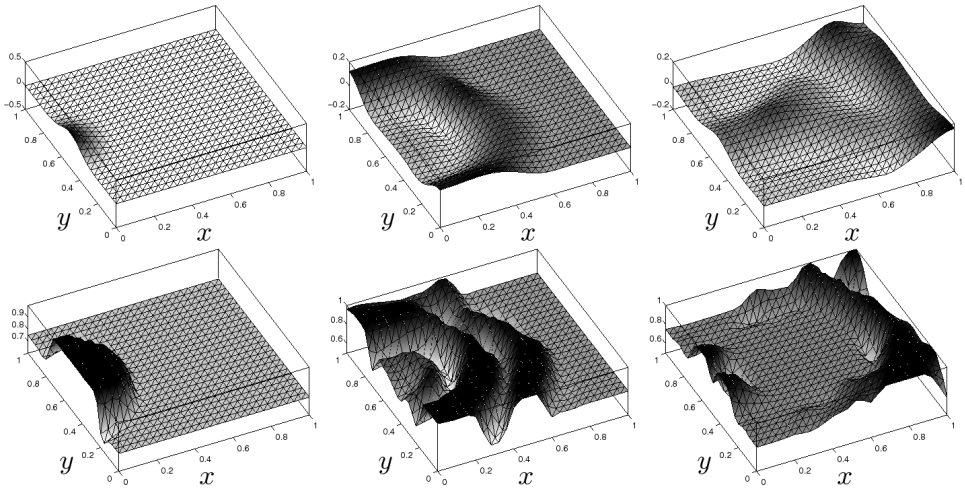

FIGURE 8. Measurements u^* (top) and $\mathfrak{h}'_\delta(\nabla u \cdot \nabla q)$ (bottom) for timesteps 5, 15 and 25. The data here corresponds to the top left plot in Figure 7, and u^* is interpolated onto the mesh used for the calculation of u and q.

References

[1] H. T. Banks and K. Kunisch. *Estimation techniques for distributed parameter systems*, volume 1 of *Systems & Control: Foundations & Applications*. Birkhäuser Boston Inc., Boston, MA, 1989.

[2] Emmanuel Nicholas Barron and Robert Jensen. The Pontryagin maximum principle from dynamic programming and viscosity solutions to first-order partial differential equations. *Trans. Amer. Math. Soc.*, 298(2):635–641, 1986.

[3] M. P. Bendsøe and O. Sigmund. *Topology optimization*. Springer-Verlag, Berlin, 2003. Theory, methods and applications.

[4] Michele Benzi, Gene H. Golub, and Jörg Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.

[5] George Biros and Omar Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. I. The Krylov-Schur solver. *SIAM J. Sci. Comput.*, 27(2):687–713 (electronic), 2005.

[6] Liliana Borcea. Electrical impedance tomography. *Inverse Problems*, 18(6):R99–R136, 2002.

[7] Jesper Carlsson. Pontryagin approximations for optimal design of elastic structures. *preprint*, 2006.

[8] Jesper Carlsson, Mattias Sandberg, and Anders Szepessy. Symplectic pontryagin approximations for optimal design. *preprint*, 2006.

[9] M. G. Crandall, L. C. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 282(2):487–502, 1984.

[10] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[11] FEniCS. FEniCS project. URL: urlhttp//www.fenics.org/.

[12] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[13] J. Hoffman, J. Jansson, A. Logg, and G. N. Wells. DOLFIN. URL: urlhttp//www.fenics.org/dolfin/.

[14] J.-L. Lions. *Optimal control of systems governed by partial differential equations*. Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York, 1971.

[15] M. Sandberg and A. Szepessy. Convergence rates of symplectic Pontryagin approximations in optimal control theory. *M2AN*, 40(1), 2006.

[16] Mattias Sandberg. Convergence rates for numerical approximations of an optimally controlled Ginzburg-Landau equation. *preprint*, 2006.

[17] Curtis R. Vogel. *Computational methods for inverse problems*, volume 23 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. With a foreword by H. T. Banks.

CSC, Numerical Analysis, Kungl. Tekniska Högskolan, 100 44 Stockholm, Sweden;
E-mail address: jesperc@kth.se